



Prof. Antony Unwin, Alexander Pilhöfer  
Lehrstuhl für Rechnerorientierte Statistik und Datenanalyse  
Institut für Mathematik  
Universität Augsburg  
<http://stats.math.uni-augsburg.de/>

## Statistik I

### Übungsblatt 3

**Abgabe:** Dienstag 08. Mai 2012, bis spätestens 12.00 Uhr; Briefkasten: Statistik I oder per email an die Übungsleiter

Die Aufgaben können auch in 2er-Gruppen bearbeitet und abgegeben werden!

- Erstellen Sie für den *Irish Exams* Datensatz in R einen Spineplot der Variablen *Field* ( $x$ ) und *Sex* ( $y$ )! **(1P)**
  - Sortieren Sie die Kategorien von *Field* im Plot nach dem Anteil der Frauen! **(1P)**
  - Berechnen Sie die Korrelationen zwischen den Fächern!  
(Verwenden Sie `cor` und `use='pairwise.complete'`) **(1P)**
  - Für welche Fächerkombinationen unterscheiden sich Männer und Frauen hinsichtlich der Korrelation am meisten? **(1P)**

- Schätzer (5P)** Seien  $X_1, X_2, \dots, X_n$  unabhängige, identisch verteilte Zufallsvariablen, deren Verteilung den Erwartungswert  $\mu$  und die Varianz  $\sigma^2$  besitzt.

Bearbeiten Sie die nachfolgenden Teilaufgaben, indem Sie hierbei jeden Schritt angeben und begründen.

- Zeigen Sie, dass das arithmetische Mittel  $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$  ein erwartungstreuer Schätzer für  $\mu$  ist. Berechnen Sie den mittleren quadratischen Fehler dieses Schätzers.
- Zeigen Sie, dass die Stichprobenvarianz  $\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  ein erwartungstreuer Schätzer für  $\sigma^2$  ist. Berechnen Sie den Bias des Schätzers  $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  für  $\sigma^2$ .
- Ist  $\mu$  bekannt, so ist  $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$  ein erwartungstreuer Schätzer für  $\sigma^2$ .

- Umfrage (5P)**

In einer aktuellen Umfrage haben 610 von 1017 Befragten angegeben, dass sie glauben, dass die Piratenpartei 2013 in den nächsten Bundestag einziehen wird.

Quelle: <http://www.zeit.de/politik/deutschland/2012-04/yougov-piraten-umfrage>

- Diskutieren Sie, ob obiger Sachverhalt wie folgt formal gefasst werden darf: Eine Zufallsstichprobe  $x_1, x_2, \dots, x_n$ —unabhängig, identisch verteilt—stammt aus einer Bernoulli-Verteilung ( $x \in \{0, 1\}$ ):

$$f(x; p) = p^x q^{(1-x)} \quad p + q = 1, p, q > 0.$$

Spezifizieren Sie hierbei explizit die Interpretationen aller mathematischen Größen des obigen Modells im Anwendungskontext der Umfrage.

- Welchen Schätzer  $\hat{p}$  würden Sie intuitiv für  $p$  vorschlagen?
- Leiten Sie den Maximum-Likelihood-Schätzer für  $p$  im offenen Intervall  $(0, 1)$  ab. Für welche Umfragesituationen würde dieser Schätzer im Intervall  $(0, 1)$  nicht existieren?

- (d) Plotten Sie in R die entsprechende Likelihood-Funktion und machen Sie sich ein Bild über den Maximum-Likelihood-Schätzer.

#### 4. Umfrage II (5P)

Betrachtet wird die Online-Umfrage aus Aufgabe 3.

- (a) Bestimmen Sie nach den drei in der Vorlesung besprochenen Konfidenzschätzverfahren

- 'normal approximativ',
- 'normal plug-in',
- 'normal genau'

konkrete Schätzintervalle für den Proportionsparameter  $p$  der Binomialverteilung aus dem Umfragebeispiel zum Konfidenzniveau 0.99. Diskutieren Sie Ihre Ergebnisse.

- (b) Simulieren Sie in R die obige Umfrage 1000 Mal und berechnen Sie jeweils die Schätzung für den Parameter  $p$  und die obigen Konfidenzintervalle. Wie oft liegt der Parameter innerhalb der dieser Intervalle?

#### 5. Almighty Dollar (5P)

Betrachten Sie die Grafik unter folgendem Link:

<http://awesome.good.is/transparency/web/1002/almighty-dollar/flat.html>

- Was genau wird dargestellt?
- Welche Kernaussage soll die Graphik Ihrer Meinung nach vermitteln?
- Beurteilen Sie die Graphik in Bezug auf Übersichtlichkeit, Ästhetik, Suggestivität.
- Welche Aussagen lassen sich auf Basis der Grafik treffen und welche nicht?
- Welche Informationen würden Sie hinzuziehen wollen, um statistische Aussagen über den eventuellen Zusammenhang zwischen Glaube und Einkommen treffen zu können?